Digital divide across the European Union

Frederico Cruz-Jesus *, Tiago Oliveira, Fernando Bacao

ISEGI, Universidade Nova de Lisboa 1070-312 Lisbon, Portugal

ARTICLE INFO

ABSTRACT

Our research analyses the digital divide within the European Union 27 between the years of 2008 and 2010. To accomplish this we use multivariate statistical methods, more specifically factor and cluster analysis, to address the European digital disparities. Our results lead to an identification of two latent dimensions and five groups of countries. We conclude that a digital gap does, in fact, exist within the European Union. The process of European integration and the economic wealth emerge as explanatory factors for this divide. On the other hand, the educational attendance is not proven to be significant, as one would expect.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

The attention given by leaders from all over the world to the concept of information society and the potential for a digital divide has risen significantly in recent years. At the World Summit on the information society, it was declared that the global challenge for the new millennium is to build a society "*where everyone can create, access, utilize and share information and knowledge, enabling individuals, communities and peoples to achieve their full potential in promoting their sustainable development and improving their quality of life*" [1,2].

The European Union (EU) has just released the Europe 2020 Strategy, which seeks to lead to "*a smart, sustainable and inclusive growth for European Economy*" [3] and "*to exit the crisis and prepare the EU economy for the challenges of the next decade*" [4]. This economic growth will be accomplished by (among other things) developing a (digital) economy based on knowledge and innovation [3]. The Digital Agenda for Europe is included in the Europe 2020 Strategy as one of the seven strategy flagships. It aims to define the central role that the use of information and communication technologies (ICT) must play if Europe wishes to realize its ambitions for 2020 [4]. The European Commission has, recently, also earmarked 1 billion Euros extra to "*help rural areas get online,*

bring new jobs and help businesses grow" [5]. It is predicted that the ICTs "*will help create around 1 million jobs in Europe and a broadband-related growth of economic activity of 850 billion Euros between 2006 and 2015*" [5]. Also during the period between the year 2000 and 2006, it is estimated that the European structural funds spent 5.5 billion Euros on information society boosting programmes [6]. Therefore, considering the importance of the digital development to the EU, expressed by these measures, digital inequalities must be detected and corrected in order to avoid jeopardizing the objectives of the Europe 2020. Thus it has become fundamental to know what the current situation is regarding the digital divide within the whole EU. The best way to assess this is by studying the national realities regarding the multiple dimensions of the digital development, for a posterior comparison of each country with the rest of the European ones.

Keeping this goal in mind, although several authors have focused on understanding and measuring the digital divide, there is a lack of studies capable of taking a wide snapshot of the EU reality with a wide set of relevant and updated indicators. Considering the importance that the European Commission gives to a homogeneous digital development amongst all of its members, the first step to take toward this development is to assess the current situation within the Union. The current research helps to do this and sheds light on the issue in order that efficient policies may be deployed. We therefore intend to provide a complete and updated analysis of digital asymmetries within the 27 Member States of the European Union (EU-27), with data pertaining to the year 2010, and answer the following questions: (1) What are the most significant latent dimensions of the European digital divide,

* Corresponding author at: Rua Doutor Rafael Duque n°9 6°dto, 1500-249 Lisboa, Portugal. Tel.: +351 914 162 677.
E-mail addresses: frederico.mcruz@gmail.com (F. Cruz-Jesus), toliveira@isegi.unl.pt (T. Oliveira), bacao@isegi.unl.pt (F. Bacao).

so that countries may act on them?; (2) Is the digital divide a consequence of economic wealth and educational attendance, as suggested by some authors [6–8]?; (3) Is the European digital divide widening or narrowing?; In answering these questions, the paper is organized as follows: Section 2 presents a literature review of digital divide and digital development; in Section 3 we develop a theoretical framework for measuring the digital divide; Section 4 includes the analyses undertaken on the data collected, using factor and cluster analysis, and comparison with the previous two years (2008 and 2009); Section 5 presents the discussion of findings, while Section 6 presents the conclusions.

2. Digital divide and digital development

Even today the mystery remains about who the first person was to use the term digital divide, and when this happened [9]. However, in the literature and forums about the subject, it is widely stated that the term was coined in the mid-1990s by the former Assistant Secretary for Communications and Information of the United States (US) Department of Commerce, Larry Irving Junior [10]. According to his own report, he used the term to describe the social division between those who were very involved in technology and those who were not. Later on he stated that the term digital divide was "*appropriated from an unknown source and redefined by the US Department of Commerce in the process of preparing the third 'Falling Through the Net' report*" [9].

The digital divide was initially understood in a binary way, meaning that there was a choice between "*has*" and "*has not*" access to ICT. While useful for describing the limits of various social and technological inequalities, this binary classification was very reductive, imprecise, and inaccurate. Consequently, the subject evolved from the binary understanding between "*has*" versus "*has not*" to focus on the reasons why disparities in access and use really existed. As a consequence, it was discovered that geographic area was an important factor in defining the divide between information haves and have-nots. Individuals belonging to ethnic minorities, or with lower incomes, were also more vulnerable to asymmetries in the access to digital technologies [9]. In other words, the very same term digital divide underwent considerable evolution as the subject started to be understood as a multifaceted issue. Hence, it is widely recognized today that the initial binary definition was narrow, since other factors need to be considered [11]. A widely accepted and repeated definition of digital divide is the one provided by the Organization for Economic Co-operation and Development (OECD): "*the term digital divide refers to the gap between individuals, households, businesses and geographic areas at different socio-economic levels with regard both to their opportunities to access ICT and to their use of the Internet for a wide variety of activities*" [12]. Therefore, digital divide is today understood to be a complex, multidimensional phenomenon [13–16].There are two types of digital divides. The first is located at an international level, that is, between different countries. The second is located at an intra-national level, or within a country. In both types of digital divide gaps can occur regarding access to ICT between regions, or groups of individuals, when characteristics of different nature exist [17,18]. Some authors have demonstrated that the domestic digital divide is characterized by a higher risk of digital exclusion of the elderly, women, population with lower income, education attainment, those with disabilities, those living in rural areas, and ethnic minorities [19–28]. Hsieh and Rai [13] showed that economically advantaged and disadvantaged people also have very different post-implementation behaviour regarding the use of ICT. These authors concluded that economically advantaged people have a "*higher tendency to respond to network exposure*", using these technologies with much more confidence than the disadvantaged. They named these inequalities about access and use of ICT as "first

order" and "second order" digital divides, respectively, adding greater complexity to the phenomena. Moreover, according to Dewan and Riggins [29], digital disparities may also be found at an organizational level, in which "*large organizations are more likely to adopt innovations and advanced ICT solutions than smaller organizations*". Hence, the digital divide can represent a threat to all of the *e-strategies* around the world, including the Digital Agenda for Europe [30,31].

Development and use of ICT have undergone exponential growth in recent decades. These technologies are playing a decisive role in improving almost every aspect of our societies [32], including business transactions, communications, economics, and politics [33]. Wattai and Schuff [34] studied the impact of the Web 2.0, especially the influence of social networks on politics, in the 2008 US Primary Presidential Campaign, and concluded that the Internet is changing the very nature of political competition. Carlsson [35] studied the effects of the ICT in the economy, comparing the potential of these technologies to the so-called "*general-purpose technologies* (*GPT*) *which in the past revolutionized the economy*", such as the transportation and communications technologies in the 19th century, the Corliss steam engine, or the electric motor. He concluded that ICT appears to have an even greater impact on the economy since "*it affects the service industries* (e.g. *health care*, *government*, and *financial services*) *even more profoundly than the goods-producing industries, and these service sectors represent over 75% of GDP*". Jalava and Pohjola [36] showed that the ICT contribution to Finland's GDP between 1990 and 2004 was three times greater than the contribution of electricity from 1920 to 1938. Moreover, new types of interactions, or advanced services, are becoming more and more common. These include e-commerce, e-government, e-health, e-learning, e-banking, e-finance, and others [6,23,37–39]. Actions and technologies like Internet surfing, YouTube, social networking, on-line job seeking, email, wiki-sites, and access to online libraries are gaining room in our daily routines, improving the way people interact with each other. These factors are drawing strong distinctions between individuals who have access to privileged information and those who have not [40]. The emergence of ICT is even changing the notion of literacy, considering that the inability to use these technologies is creating an entirely new group of disadvantaged people who were considered "literate" in the past [41]. Therefore, there is evidence that ICT positively affects the economy and welfare in several dimensions [37,42]. ICT creates competitive advantages in enterprises, improves national health systems [43] through e-health, improves educational systems [13,44] through e-learning, which creates new opportunities, all of which reduces distance constraints and creates new industries that generate new employment opportunities [45,46]. Thus, for these benefits to be realized, certain obstacles need to be overcome, especially inequalities both between and within countries regarding the access to these technologies.

3. Measuring the European digital divide

3.1. Framework

Due to ICT's importance in the improvement of the economy and social care, the problem of how to measure the digital divide has gained importance in terms of research [30,31,42,47,48]. Wang et al. [49] identified 852 journal articles and books published between 2000 and 2009, with more than 26,000 citations using the term "digital divide" as keyword, in order to map the intellectual structure on this subject. However, despite the increasing attention that this phenomenon has received, measuring the access and diffusion of ICT is a complex task plagued by several constraints. Firstly, there is no single and standardized definition of

digital development, information society, or digital divide [47]. As a result, considerations about these subjects differ between countries, geographical areas, organizations, and models of information society [45]. In fact there are several models of information society, such as those from Finland, Singapore, and the US (Silicon Valley). Each emphasizes specific characteristics and objectives which are in line with their own respective national realities [45]. The second constraint is related to a lack of harmonized data available when considering the analysis for multiple countries. Hence, there is usually a "trade-off" between the depth and the width of the analysis. This means that the more indicators that researchers try to use, the fewer are the countries that can be included in the analysis [31]. This constraint affects accurate attempts to measure digital inequalities particularly, in a very significant way: if the data collected are scarce in terms of number of individuals (countries) the conclusions will be limited. On the other hand, if the data are scarce in terms of number of variables, the conclusions may be misleading, considering that "*sometimes digital divide indicators take different or even contradictory values*" [47], i.e. a country may have higher levels of broadband penetration rates but, at the same time, lower levels in Internet adoption by individuals and firms than another country. Therefore, if it is assumed that the broadband penetration rate is synonymous with Internet adoption (which seems at a first sight fair enough), the conclusions of the analysis will be misleading. Because of this, a consistent theoretical framework should be developed prior to the use of data, especially if this is scarce in terms of variety of indicators.

For these reasons, the use of indices to measure the digital divide has emerged as relatively popular alternatives to the use of ICT-related indicators. Various compound ICT-related indicator measures have appeared, such as, for example, the Digital Access Index (ITU, 2003), the Digital Opportunity Index (ITU, 2005), the Technology Achievement Index (UNDP, 2001), the Information Society Index (IDC, 2001), and the Infostate, which is an overall set of indicators comprising two components, info-density and info-use [50]. There are even authors who have proposed new indices based on others, in a attempt to achieve a standardized solution for measuring ICT adoption and use (see for example [51]). The use of indices can summarize complex and multidimensional phenomena such as the economic or digital development. They are also easy to interpret, and allow for easier tracks of changes across time periods [52]. However, they also entail several constraints [53]. One of the main arguments against their use rests on the oversimplification of complex interrelations, i.e. the indices reduce the digital divide to a single value, which may be misleading in some instances [47]. Another limitation is related to the variables that are included in each index [18,52]. For instance, besides technologies that have already reached some level of widespread use (e.g. TV, fixed telephone lines, PCs, and the Internet), it is fundamental to measure new emerging types of ICT (e.g. broadband, Internet access via mobile devices, e-banking services, among many others). New technologies should be continuously incorporated into the indices. Note that several of the ones mentioned above are almost ten years old, and some are even older. The advent of new types of ICT is very dynamic. Ten years ago the use of Internet via mobile devices, the use of e-banking, e-government, among other electronic services, were, at the very best, almost negligible compared with to what it is now. Finally, another limitation of these compound measures is related to the weight that each indicator (variable) has in calculating the index "*which could be the subject of political dispute*" and therefore not reliable for a specific subset of countries or ICT dimensions [52,53].

Due to these limitations and the fact that we wish to assess the European digital divide, explaining the specific latent dimensions within the Union, and group the countries into similar digital profiles (clusters), we find the use of multivariate methods more suitable and reliable. Several authors in the past have also opted to follow this methodology (see, e.g. [31,37,54]). Moreover, considering that the EU has a unified statistics system, the Eurostat, the problem of data availability described above is mitigated by the fact that Eurostat has the data needed, and those data are fairly well harmonized.

In order to collect the ICT-indicators to measure the digital development of countries within EU, we conceptualise the ICTs as general purpose technologies (GPTs). GPTs are characterized by technological innovations that have the potential to positively affect multiple industries and society sectors [55], as is the case of ICT [35]. We must therefore face the digital divide as a multidimensional issue. To do this we use multiple indicators to represent the ICT development of a country. According to the recommendations of the OECD [30], the variables that should be used to measure the digital divide vary with the goals of the research. For instance, if we wish to measure the internal or domestic digital divide we should "drill down" the ICT level indicators by groups such as gender, age, income, education, geographical place, and so on, which are more likely to present disparities between categories. To measure the digital divide between countries, the indicators should refer to the aggregated national reality. Because our goal is to investigate the divide within the EU, we will follow the second recommendation. A limitation for this option must be recognized, however. The variables will not reflect any type of internal disparities (domestic digital divides) in each country within itself. It should be kept in mind that these potential domestic digital divides are more likely to occur in bigger (thus more heterogeneous) countries, as well as in countries of lower economic development. Countries with these characteristics are potentially more likely to reveal domestic digital disparities because smaller and richer societies are easier to connect than those which are bigger and poorer [51,56].

Recent studies have suggested that the international digital divide is mainly a consequence of economic inequalities between countries. The terms "information rich" and "information poor" have appeared to classify countries in terms of their digital development. Besides economic development, countries with lower educational attainment also tend to present lower rates in the use and adoption of ICT [57–62]. Dewan and Ganley [57] showed that developing countries are slower to achieve digital development, but by focusing on certain technologies, particularly the availability of PCs and Internet with a cost-reduction policy, the cross-technology diffusion effects of these combined technologies will help to appreciably narrow the divide.

3.2. Data

Consistent with our theoretical framework, which comprises ICT into the category of GPT, for measuring the levels of digital development across the EU-27, we used 16 variables that are compatible with recommendations from the OECD and European Commission. These indicators were selected by combining a mix of earlier studies with some recommendations from the organizations mentioned. We were thereby able to obtain 15 of 16 variables used in our analysis from the Information Society Statistics Category in the Eurostat website–all pertaining to the year 2010. Data for the 16th variable were obtained from the World Bank database. The fact that all variables were obtained from official entities, and pertain to the year 2010, guarantees that the results of the analysis enjoy a high degree of reliability. Table 1 shows the variables collected.

As in other studies focusing on the digital divide, we sought to include indicators that measure the ICT infrastructure of each

Table 1
Acronyms, descriptions and literature support of variables.

Code	Variable	Support
HsInt	Percentage of households having access to the Internet at home	[11,31,37,54,58]
BroRt	Broadband penetration rate	[11,31,37,54,63]
IntPop	Percentage of population regularly using the Internet	[37,54,63]
Mob	Percentage of population using mobile devices to access the Internet	[14,25,26,64]
IntSrc	Percentage of population using Internet for finding commercial information	[37,54,65]
Cost	Percentage of households without Internet because of the access costs	[31,41,57,60]
eBank	Percentage of population using e-banking services	[4]
eLearn	Percentage of population using e-learning services	[4,37]
email	Percentage of population using e-mail	[8,65,66]
eHealth	Percentage of population using Internet for seeking health information	[4]
eGovI	Percentage of population using Internet for interaction with public authorities	[4,37]
eGovE	Percentage of enterprises using Internet for interaction with public authorities	[4]
eGovS	Percentage of government services available online	[4,31]
eCom	Percentage of enterprises selling online	[31]
Serv	Number of secure servers per million inhabitants	[11,31,63]
eSafeE	Enterprises having a formally defined ICT security policy with a plan of regular review	[67–69]

country along with their pervasiveness at multiple units of adoption (individuals, and enterprises).

The percentage of households connected to the Internet (HsInt), and the broadband penetration rate (BroRt) are often used in the literature to measure the digital development/divide [11,31,37,54,58]. These variables express the connectivity level in terms of ICT infrastructure. The Internet secure servers (Serv) are also a specific ICT infrastructure of e-commerce, allowing secure electronic business transactions [31]. Consistent with the literature [41,57,60], the Internet costs were also considered in our analysis, since this is strongly negatively correlated with the digital development. For this purpose we included the percentage of households without Internet because of the access costs (Cost). Web browsing and using e-mail are probably some of the most general and popular actions that individuals can practice through the use of ICT [65]. Hence, the percentage of population regularly using the Internet (IntPop), and the percentage of population using e-mail (email), is an effective way for assessing the use of ICT of individuals for general purposes. The recent emergence of the mobile phones that allow Internet access is also an important aspect of the information society [14]. However, as Kauffman and Techatassanasoontorn posit, "*little attention has been devoted to evaluate the extent of the divide and empirically examine determinants of the diffusion of digital wireless phones across countries*" [64]. For this reason, we find it particularly interesting to measure the percentage of population using mobile devices to access the Internet (Mob). The percentage of population using Internet for finding commercial information (IntSrc) allows us to assess the role that the Internet has in commerce (on the individual side). Considering that we are attempting to analyse the digital divide at a European level, we also find it particularly relevant to consider the position of the European entities concerning the digital divide. The European Commission, via the Digital Agenda for Europe [4], emphasizes the role of specific electronic services, more specifically e-health, e-learning, e-banking, and e-government. E-banking, e-learning, and e-health are considered to be "*some of the most innovative and advanced online services*" [4]; e-government services are also highlighted in the Digital Agenda, since "*despite a high level of availability of e-government services in Europe, differences still exist amongst Member States*" [4]. The inclusion of these indicators, related to how the use of these advanced services also allows us to analyse the so called "second order digital divide", expands our focus from mere ICT adoption, to include the manner in which it is used by individuals [13]. For these reasons the percentage of population using e-health, e-learning, e-banking, and e-government services (eHealth, eLearn, eBank, eGovI, and eGovS, respectively) were also included in our analysis. As mentioned we also took into account the pervasiveness of ICT in

adoption units other than the individuals. Hence, the percentage of enterprises selling online (eCom) is an important indicator of electronic commerce. Likewise, at the individual level e-government is also present in the relationship between firms and public authorities. We thus include the percentage of enterprises using Internet for interaction with public authorities (eGovE). Finally, consistent with the literature, security is recognized as a major issue in the ICT by enterprises [68]. Moreover, the awareness of enterprises to the importance of electronic security is a fundamental requirement for the pervasiveness of electronic business transactions [67,69]. For these reasons, we also include the percentage of enterprises having a formally defined ICT security policy (eSafeE). The last column of Table 1 presents the literature support (either theoretical or empirical) of the suitability of our variables to measure the digital development.

It is important to keep in mind that we do not claim that our set of variables are exhaustive in terms of expressing the digital development, but we do believe that considering the limitations of data availability, they can efficiently measure each of the EU-27 member state's digital development to a significant and wide extent.

The data used (see Table 2) show high disparities within the EU-27 related to the ICT: in Bulgaria only 2% of the population uses e-banking services, while in the Netherlands 77% do so. In Romania only 7% of the population uses the Internet for interacting with public authorities (eGovI), while in the Denmark this figure stands at around 72%, which is ten times higher. When analysing the e-government supply availability (eGovS), there are six countries with 100% services available online and nine countries with values less to 75%. Also, in the percentage of households without Internet because of the access costs (Cost), we have four countries with a value above 15% and nine under 5%. We also notice extreme asymmetries in the overall profile of the 27 countries. The Netherlands is the best-ranked country for eight of the 16 variables used, while Bulgaria and Romania are the poorest-ranked in eight of them. These uneven distributions can tell us a great deal about the asymmetries that exist between countries. Nevertheless, the dimensionality of the data used -16- makes it impossible to address the digital divide with simple univariate statistics. This is why the use of multivariate statistical methods is much more appropriate in analysing all of these digital asymmetries.

4. Methodology

4.1. Factor analysis

Factor analysis uses the correlation between variables in order to find latent factors within them [70]. In order to apply factor analysis successfully some assumptions need to be confirmed.

Table 2
Data used.

Country		HsInt	BroRt	IntPop	Mob	IntSrc	Cost	eBank	eLearn	email	eHealth	eGovI	eGovE	eGovS	eCom	Serv	eSafeE
Austria	Au	73	24	70	37	58	4	38	35	66	37	39	75	100	14	857	24
Belgium	Be	73	30	75	23	62	6	51	39	72	37	32	77	79	26	490	29
Bulgaria	Bu	33	14	42	8	26	9	2	13	35	13	15	64	70	4	73	7
Cyprus	Cy	54	23	50	10	47	9	17	23	41	21	22	74	55	7	1051	37
Czech Rep.	CR	61	20	58	9	53	5	23	22	59	21	17	89	74	20	318	21
Denmark	De	86	38	86	47	78	1	71	64	83	52	72	92	95	25	1866	43
Estonia	Es	68	26	71	35	61	17	65	33	63	35	48	80	94	10	434	11
Finland	Fi	81	29	83	26	74	5	76	70	77	57	58	96	95	16	1246	37
France	Fr	74	31	75	29	65	10	53	53	72	36	37	78	85	12	306	22
Germany	Ge	82	31	74	34	72	5	43	38	72	48	37	67	95	22	874	27
Greece	Gr	46	19	41	7	36	5	6	28	32	22	13	77	48	9	124	39
Hungary	Hu	60	20	61	14	55	16	19	33	58	41	28	71	66	8	166	9
Ireland	Ir	72	23	63	37	57	3	34	44	58	27	27	87	100	21	1005	28
Italy	It	59	21	48	25	35	4	18	38	43	23	17	84	100	4	154	29
Latvia	La	60	19	62	19	57	22	47	42	55	32	31	72	93	6	173	15
Lithuania	Li	61	20	58	13	48	8	37	25	49	31	22	95	72	22	176	25
Luxembourg	Lu	90	33	86	54	78	0	56	72	83	58	55	90	72	14	1413	28
Malta	Ma	70	29	60	20	52	1	38	43	54	34	28	77	100	16	1365	30
Netherlands	Ne	91	39	88	33	82	0	77	38	87	50	59	95	95	22	2276	29
Poland	Po	63	15	55	20	39	8	25	35	48	25	21	89	79	8	211	11
Portugal	Pt	54	19	47	19	44	8	19	42	45	30	23	75	100	19	174	22
Romania	Ro	42	14	34	5	26	22	3	20	31	19	7	50	60	6	40	9
Slovakia	Sk	67	16	73	32	62	6	33	27	70	35	35	88	63	7	128	35
Slovenia	Sn	68	24	65	24	57	13	29	47	58	43	40	88	95	10	301	16
Spain	Sp	59	23	58	24	54	10	27	39	55	34	32	67	95	12	233	33
Sweden	Sw	88	32	88	44	82	3	75	50	84	40	62	90	100	24	1266	46
U.K.	UK	80	31	80	38	63	4	45	42	74	32	40	67	98	15	1396	29
Minimum		33	14	34	5	26	0	2	13	31	13	7	50	48	4	40	7
Maximum		91	39	88	54	82	22	77	72	87	58	72	96	100	26	2276	46
Average		67	24	65	25	56	8	38	39	60	35	34	80	84	14	671	26
Std. Deviation		15	7	15	13	16	6	22	14	16	12	16	11	16	7	623	11

Using this technique depends on the correlation structure within the input data [71]. Hence, we need to confirm that this correlation exist, otherwise the factor analysis may provide weak results. Our analysis involved several steps. The first was to analyse the correlation structure of the data by using the correlation matrix. The second was to confirm the suitability of the data using the Kaiser–Mayer–Olkin (KMO). In the third step we chose the extraction method to be used. In the fourth step the number of factors to be extracted was defined and we proceeded to the interpretation of the factors based on its loadings.

The correlation matrix (see Table 3) shows that each variable has, at least, one absolute correlation coefficient of 0.55 with another variable. Although this correlation value is moderate, it ensures that all of the variables are measuring the same phenomena. We notice that some pairs of variables present extreme correlation levels. The percentage of population regularly using the Internet (IntPop) has a correlation level of 0.99 with the percentage of population regularly using e-mail (email) and a value of 0.97 with the percentage of population using the Internet for finding commercial information about products or services (IntSrc). At the other end of the spectrum we have the percentage of e-government services available online (eGovS) with a correlation level of 0.13 with the percentage of enterprises that have adopted a regular ICT security plan (eSafeE). We also notice that the same availability of e-government services online shows a low correlation (0.2) with the adoption of these services by the enterprises (eGovE). Hence, it may be that factors other than the availability of these services influence the adoption decisions. Lee et al. [72] showed that the willingness by business users to adopt public services online is significantly related to the perceived quality of those *services vis-à-vis* traditional (offline) channels. These authors showed that businesses tend to have other drivers that influence the decision of use e-government services than its

mere availability, a belief that our correlation matrix appears to support.

To confirm the suitability of the data for factor analysis, KMO was performed. It returned the value of 0.77, which expresses a good suitability [73].

As extraction method we applied the factor analysis, which is the method most widely used in Marketing and the Social Sciences [74]. Since our aim is to reduce the complexity of the problem, we had to decide how many factors we would extract from the factor analysis. There are no definitive criteria to define the number of factors to retain, but it is important to note that the decision should depend on the context of the analysis. There are three main criteria for defining the number of factors to retain; Pearson's, Kaiser's, and the Scree Plots. All of these methods were taken into consideration [74], and all yielded the same solution: the optimal number of factors to be extracted is two. As shown in Table 4, the percent of variance retained in these two factors is 76%.

Since our objective is to reduce the complexity of the data about the digital divide, in our factor analysis we used the rotation of the factors in order to achieve a better split of the original indicators in only one factor. Although there are several types of rotation, including orthogonal and oblique methods, the orthogonal ones seem to be the most widely used [75]. In particular, we applied the Varimax rotation. Varimax and Quartimax rotations led to similar results, in fact, which support the belief that our solution is based on a well explained factor structure.

To measure the scale reliability of each factor, Cronbach's Alpha was also calculated. It measures the internal consistency of each factor within itself. Nunnally [65] suggests that a value over 0.7 is considered good. The values returned were 0.97 for factor 1 and 0.85 for factor 2, which confirm the high reliability of the two factors extracted.

Table 3
Correlation matrix.

	HsInt	BroRt	IntPop	Mob	IntSrc	Cost	eBank	eLearn	email	eHealth	eGovI	eGovE	eGovS	eCom	Serv	eSafeE
HsInt	1	0.87**	0.95**	0.85**	0.94**	−0.55**	0.88**	0.74**	0.94**	0.84**	0.86**	0.55**	0.53**	0.64**	0.77**	0.49**
BroRt		1	0.83**	0.71**	0.85**	−0.55**	0.82**	0.69**	0.83**	0.75**	0.82**	0.37*	0.48**	0.65**	0.84**	0.53**
IntPop			1	0.83**	0.97**	−0.46**	0.92**	0.69**	0.99**	0.84**	0.92**	0.53**	0.45**	0.57**	0.71**	0.44*
Mob				1	0.79**	−0.49**	0.73**	0.71**	0.82**	0.70**	0.82**	0.40*	0.56**	0.43*	0.63**	0.41*
IntSrc					1	−0.43*	0.90**	0.70**	0.97**	0.87**	0.91**	0.51**	0.43*	0.62**	0.72**	0.50**
Cost						1	−0.34*	−0.37*	−0.47**	−0.33*	−0.38*	−0.54**	−0.25	−0.56**	−0.64**	−0.67**
eBank							1	0.70**	0.89**	0.78**	0.91**	0.55**	0.53**	0.58**	0.70**	0.42*
eLearn								1	0.68**	0.80**	0.75**	0.45**	0.50**	0.37*	0.54**	0.42*
email									1	0.83**	0.90**	0.50**	0.46**	0.60**	0.70**	0.42*
eHealth										1	0.86**	0.46**	0.38*	0.47**	0.61**	0.35*
eGovI											1	0.51**	0.49**	0.49**	0.74**	0.46**
eGovE												1	0.20	0.42*	0.38*	0.41*
eGovS													1	0.40*	0.39*	0.13
eCom														1	0.56**	0.47**
Serv															1	0.54**
eSafeE																1

* Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).

As mentioned above, the final step of our factor analysis is to interpret the factors extracted based on their loadings, i.e. based on the variables that contribute the most to each dimension.

Apparently the digital development can be explained by two latent dimensions, in which asymmetries between countries may, or may not, exist. The first is the *ICT Infrastructure and adoption by Population*, which is related to the availability of ICT infrastructures and their use by the population. This dimension includes the Internet and broadband penetration rates, the usage of mobile devices to access the Internet, the availability of e-government services by the supply (public) side, the adoption of e-government services by the users' (population) side, as well as the nature and intensity of Internet use. The second dimension is related to the commercial use of the ICT and its access costs and is therefore named *e-business and Internet access costs*. This dimension is related to the diffusion of e-business, including the diffusion of e-commerce, e-safety concerns by firms, and e-government, as well as the Internet access costs. We computed the factor scores for each country, and plotted it for a comparison analysis (see Fig. 1).

Thus, Denmark, the Netherlands, and Sweden are the best-ranked countries for the two dimensions extracted together. These

Table 4
Results of factor analysis and Cronbach's alpha.

Rotated factor model: Varimax		
	Factor 1	Factor 2
eGovI	**0.90**	0.32
IntPop	**0.89**	0.37
IntSrc	**0.89**	0.39
email	**0.88**	0.37
eBank	**0.88**	0.31
eHealth	**0.88**	0.23
HsInt	**0.86**	0.46
Mob	**0.80**	0.32
eLearn	**0.77**	0.25
BroRt	**0.76**	0.49
eGovS	**0.59**	0.07
eSafeE	0.17	**0.83**
eCom	0.40	**0.68**
Serv	0.55	**0.61**
eGovE	0.34	**0.57**
Cost	−0.14	**−0.91**
Variance (%)	51%	25%
Variance total	51%	76%
Cronbach's Alpha	0.97	0.85

Note: Variables are marked according to factor loading.

North European Countries present high levels of both *ICT Infrastructure and adoption by Population* and *e-business and Internet access costs*. On the other hand we have Bulgaria and Romania as the least digitally developed countries in the EU, showing extremely low levels for both dimensions.

4.2. Cluster analysis

After the use of factor analysis – in which we found two latent dimensions on the digital divide – we used a cluster analysis to group the countries by similarity criteria, both for factors and the original 16 variables. The use of cluster analysis involves two main methods, either hierarchical or non-hierarchical. The methodology used for clustering based on factors and the original 16 variables were similar. We first ran a hierarchical procedure to define the number of clusters to extract, since in these procedures the number of clusters depends on the data, which means that we do not need to define a priori how many clusters we wish to generate. The solution based on hierarchical procedures depends on the distance measurement and the algorithm used [76]. In particular, we used Single, Centroid, Complet, and Ward's methods. Moreover, different distances were used. Euclidean distance, squared Euclidean distance, the city-block approach, and the Minkowsky distance were taken into consideration. All of these approaches returned similar results, and the solution was made based on its performance, that is, based on the analysis of the R-square and dendogram. Then, the best combination of hierarchical procedures was used to generate the initial seeds of the non-hierarchical algorithm – k-means. According to Sharma [75], this approach, tend to yield better results. Following the generation of the clusters, we classified them, and their countries, based on a "profiling analysis", in other words, on the average of each cluster for each factor/variable. Finally, we performed a Kruskal–Wallis test to verify if each variable presented statistically different values in each cluster.

4.2.1. Cluster analysis using factor scores

The solution of the hierarchical methods of the cluster analysis based on factor scores is given by the dendogram (see Fig. 2). The horizontal axis measures the distance and the vertical axis represents the countries. From left to right, the dendogram maps the clusters' formation. Thus, Ireland and Malta is the first pair of countries to form a group, which means that considering the two dimensions on the digital divide, these countries are those with digital profiles that are most similar, followed by the Netherlands
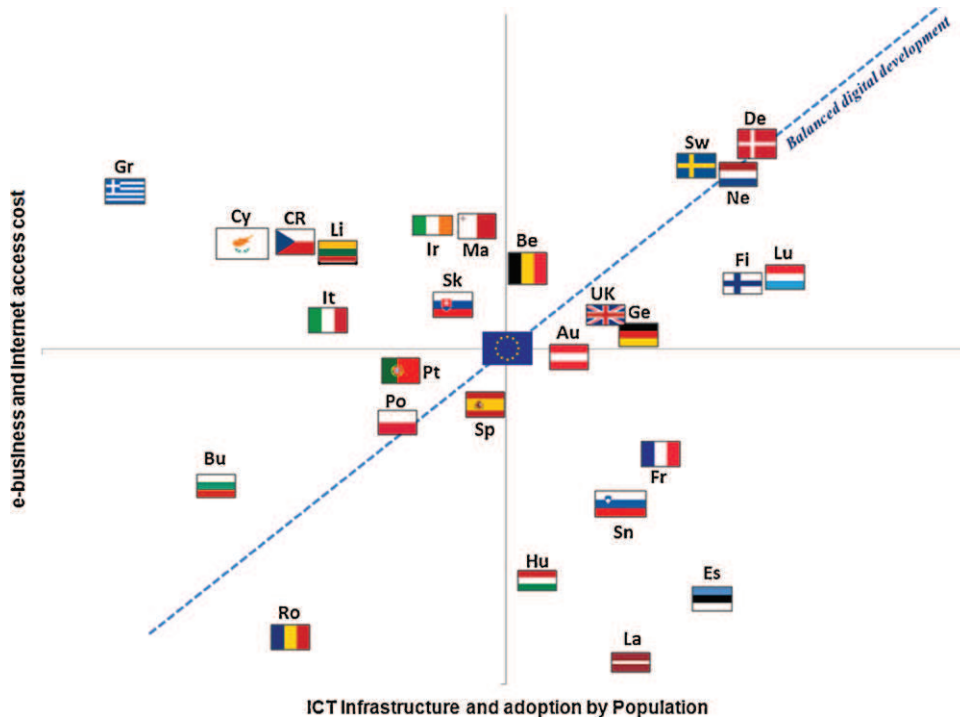
Fig. 1. Countries coordinates on extracted factors.

and Sweden, Germany and the United Kingdom, Finland and Luxembourg, and finally, the Czech Republic and Lithuania. As the algorithm continues, all countries are grouped into clusters. As mentioned, the number of clusters to extract from *k*-means, as well as the initial seeds, is obtained by hierarchical methods. We opted for a five-cluster solution with the initial seeds determined by Ward's method, since this combination is, by analysis of the dendogram and the *R*-square (respectively, Figs. 2 and 7 in Appendix A), the best solution.

From the analysis of the clusters' formation with the respective average level of each factor (see Table 5), we noticed that: Bulgaria and Romania form the group of the least digitally developed countries in the EU-27, having extremely low levels on the average of both dimensions. Besides the level of *ICT Infrastructure and adoption by Population* which is by far the lowest within the EU, the *e-business and Internet access cost* is also the lowest. Hence these countries form the "*digital laggards*" cluster. Estonia, France,

Fig. 2. Ward's dendogram for the digital divide across the EU-27 (factors).

Hungary, Latvia, and Slovenia present high levels of *ICT Infrastructure and adoption by Population* and low levels of *e-business and Internet access costs*. This cluster has a highly unbalanced digital development. Therefore, we labelled these as the "*individual-side focused*" cluster. Cyprus, the Czech Republic, Greece, Italy, and Lithuania have an unbalanced digital development as well, but with opposite values for each dimension. Hence, *ICT Infrastructure and adoption by Population* is low in these countries, but on the other hand, the levels of *e-business and Internet access costs* are high. This group is labelled as the "*firm-side and low access costs focused*" cluster. Austria, Belgium, Germany, Ireland, Malta, Poland, Portugal, Slovakia, Spain, and the United Kingdom are comprised together, making this cluster the largest. It comes as no surprise that this cluster represents the average of the EU on the two dimensions of the digital divide. Despite the absence of high levels in either of the two dimensions, the fact is that neither of the two has significant negative values either. Therefore, these countries are relatively digitally developed, with balanced levels on both dimensions. Hence they are called the "*digital followers*" cluster, considering that there is other group more advanced in terms of digital development. Finally we have Denmark, Finland, Luxembourg, the Netherlands, and Sweden together. This group comprises the most digitally developed countries in the EU. These countries present the highest levels for both dimensions of digital development. Hence they are labelled as the "*digital leaders*".

The results of the Kruskal–Wallis test also show that there are significant statistical differences in the levels of *ICT Infrastructure and adoption by Population* and *e-business and Internet access cost* for each cluster at a significance level of 1%.

Although the cluster analysis may be useful to quantify the digital asymmetries between European countries, it cannot provide explanations for why these disparities exist. Therefore, and considering the importance of understanding why there is a digital divide within the EU, based on the literature, we used some of the countries' characteristics to assess if they have explanatory power on these same differences. We found particularly interesting to consider the entrance year to the Union, the economic
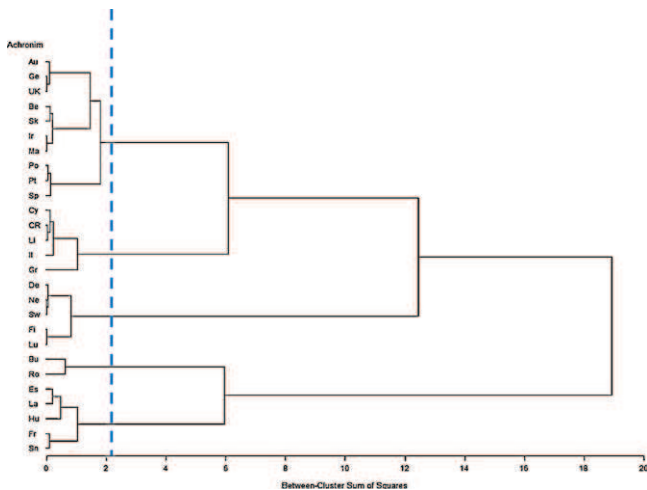
Table 5
Descriptive statistics for the identified clusters (factors).

	Digital laggards		Individual-side		Firm-side and costs		Digital followers		Digital leaders		Kruskal–Wallis
	Average	St Dev	Average	St Dev	Average	St Dev	Average	St Dev	Average	St Dev	p-Value
Factor1	−1.37	0.3	0.67	0.4	−1.30	0.5	−0.04	0.5	1.26	0.1	0.0002
Factor2	−1.42	0.7	−1.41	0.6	0.68	0.3	0.17	0.5	0.95	0.5	0.0008

wealth of each country [8,56] and also the educational attendance [8]. When considering the entrance year to the Union, the newer members are those within the "*digital laggards*" cluster. Moreover, the "*digital leaders*" cluster, includes only countries that entered the Union before the 2004 enlargement. All of the countries (10) that entered in 2004, with the exception of Poland and Slovakia, are found in clusters with uneven digital developments. We also used the Kruskal–Wallis test to verify whether the economic wealth of a country (GDP per capita in Euros) and the educational attendance (percentage of population with tertiary academic degrees) have statistically significant differences between the clusters of countries. The economic development of a country proven to be statistical significant different (p-value = 0.0024), whereas the educational attendance showed no statistical differences (p-value = 0.1865) between clusters.

4.2.2. Cluster analysis using original 16 variables

As mentioned above, the cluster analysis involved two perspectives. In the first, we used the factor scores obtained from the factor analysis to generate five groups of countries based on their digital profile similarity. In the second, the whole set of the 16 original variables was used, instead of the factor scores. Again the methodology was the same, hierarchical methods were used to define the number of clusters while non-hierarchical methods were used to obtain the final solution. Thus, the Czech Republic and Lithuania is the first pair of countries to form a group, followed by Austria and the United Kingdom, Denmark and Sweden, Ireland and Malta, and finally, Portugal and Spain. Once again, the hierarchical solution is given by the dendogram and R-square (see Figs. 4 and 8 in Appendix A, respectively). The result is a five-cluster solution obtained by using the Ward's method. As in our previous cluster analysis (based on factor scores), the number of

clusters and initial seeds were once again used as input for the k-means algorithm, which provides the final solution.

From the analysis of the averages of each cluster on the original 16 variables (see Table 6 in Appendix A), some conclusions can be drawn. Bulgaria and Romania formed the group of countries with the lowest level in 15 of the 16 variables used, showing the high asymmetry between these countries and the rest of the EU. Estonia, France, Hungary, Latvia, and Slovenia are the group of countries which is in the mid-position in 8 of the 16 variables, i.e. half of the entire dataset. It is noticeable for being the group with the second highest levels of penetration of "advanced services" like e-banking, e-learning, e-health, and e-government amongst the individuals. On the other hand, it stands out as the group with the highest Internet access costs, although the difference between it and cluster one is negligible. Cyprus, the Czech Republic, Greece, Italy, Lithuania, Poland, Portugal, Slovakia, and Spain form the cluster with the second-poorest levels for 11 of the 16 variables, more specifically in the adoption of "advanced services" by individuals and ICT infrastructure. On the other hand, these countries are relatively well positioned in some features of digital development, particularly in the business dimension and in the Internet access costs. Hence, the percentages of enterprises using the Internet for interacting with public authorities (eGovE) and having an ICT security plan (eSafeE) are the second-best ranked in the entire EU, while the percentage of households without Internet because of the access costs is less than half of the previous two clusters. Austria, Belgium, Germany, Ireland, Malta, and the United Kingdom form the second-best ranked cluster in 9 indicators, and is even in first place regarding the availability of public services online. However, these countries need to improve the penetration of the "advanced services" amongst the population in order to take advantage of their already significant infrastructure. Finally,
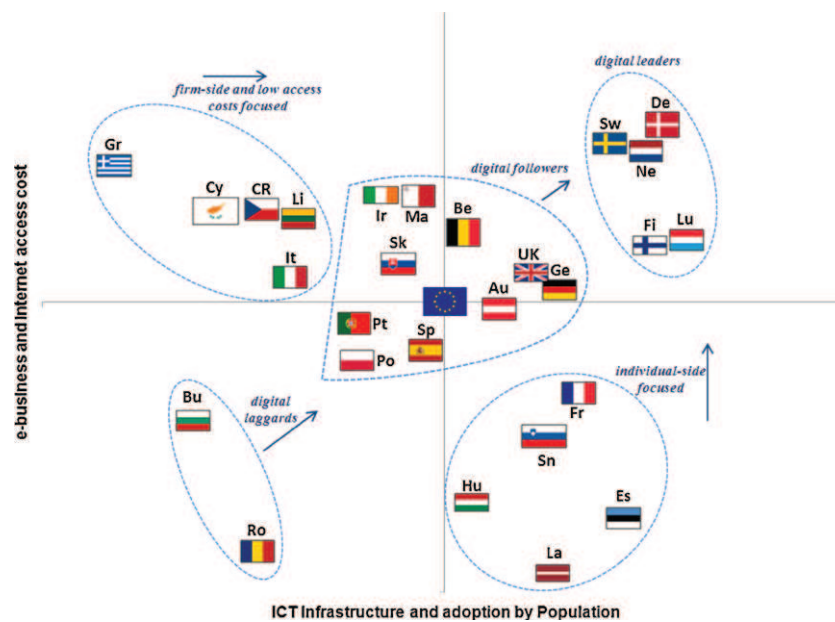

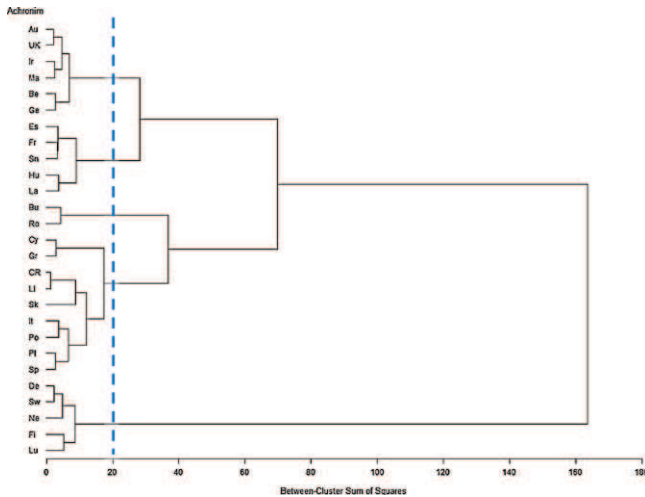**Fig. 3.** Cluster analysis on factor scores.

Fig. 4. Ward's dendogram for the digital divide across the EU-27 (variables).

Denmark, Finland, Luxembourg, the Netherlands, and Sweden are the high point of digital development in the EU-27. This group presents the highest levels in 15 of the 16 variables included.

The Kruskal–Wallis test shows that all variables except the eGovS present statistically significant differences at the 5% level for each cluster. The differences in eGovS are statistically significant only at 10%. This can be explained by the fact that the percentage of public services online can, but it should not, be independent of the diffusion of ICT within a country, since government policies can implement an e-government service only from the supply side without, however, accomplish its final objective – the actual use. The relationship between the entrance year to the EU and the cluster membership continues to appear to be correlated. The newest members continue in a cluster apart, within the "*digital laggards*". The cluster comprising the "*digital leaders*" continues to include only those that entered the EU before the 2004 enlargement. Moreover, the majority of countries that entered in 2004 are spread out amongst the "*individual-side focused*" and "*firm-side and low access costs focused*" clusters, with Malta within the "*digital followers*". Despite some positive aspects of each cluster, they all continue to show imbalances in terms of digital development. When assessing the relationship between the digital divide with the economic wealth and education attendance, once again the economic wealth came up with statistically different values across the countries (p-value < 0.001) while the education attendance presented marginal differences across the clusters (p-value = 0.1102), reason why we rejected the null hypothesis

that considers the education has an explanatory factor of the digital divide.

4.2.3. Differences between cluster analysis based on factors and original variables

When comparing the results from the cluster analysis based on factor scores and those based on the original 16 variables, some conclusions may be drawn from the results. The number of clusters present within the EU when it comes to the digital divide was the same – five. Moreover, the composition of these clusters is very similar (see Fig. 5). From the 27 European countries, only four have changed their position. When considering the clusters based on factor analysis, the "*digital laggards*", "*individual-side focused*", and "*digital leaders*" clusters maintained exactly the same composition. Only Poland, Portugal, Slovakia, and Spain moved from the "*digital followers*" to the "*firm-side and low access costs focused*". This movement is supported by an increased level of detail when using the 16 variables instead of the two latent dimensions. If in the cluster analysis based on factor scores, "*firm-side and low access costs*" cluster presented higher levels on *e-business and Internet access cost* than the "*digital followers*" cluster, using the original 16 variables for cluster analysis, this holds true only for some contributing variables of that dimension. Hence the first cluster present higher levels of enterprises selling online (eCom) and lower internet access costs (Cost), which can be explained by refinement of our analysis involving all variables that led to the disaggregation of the second dimension, providing a more detailed analysis. In this way, the cluster analysis using the original 16 variables was worth doing because it allowed us to assess the representativeness of the factor analysis in explaining the European digital divide. We can conclude, therefore, that the slight difference between the two cluster analyses reinforces the suitability of the two dimensions previously extracted.

4.3. Short-term digital divide evolution

After assessing the suitability of the two latent dimensions, confirmed by the slight differences between the two cluster analyses, the last step of our analysis of the European digital divide is to plot the path movements of each European country for the years between 2008 and 2010. For this purpose, the variables were previously normalized (as we used the correlation matrix in our factor analysis) with the averages and standard deviations of 2010. In this way, although losing scale, the values are still directly comparable across years. Following the normalization we calculated the factor scores based on the standardized coefficients. Due to limitations of data availability, we were not able to extract the

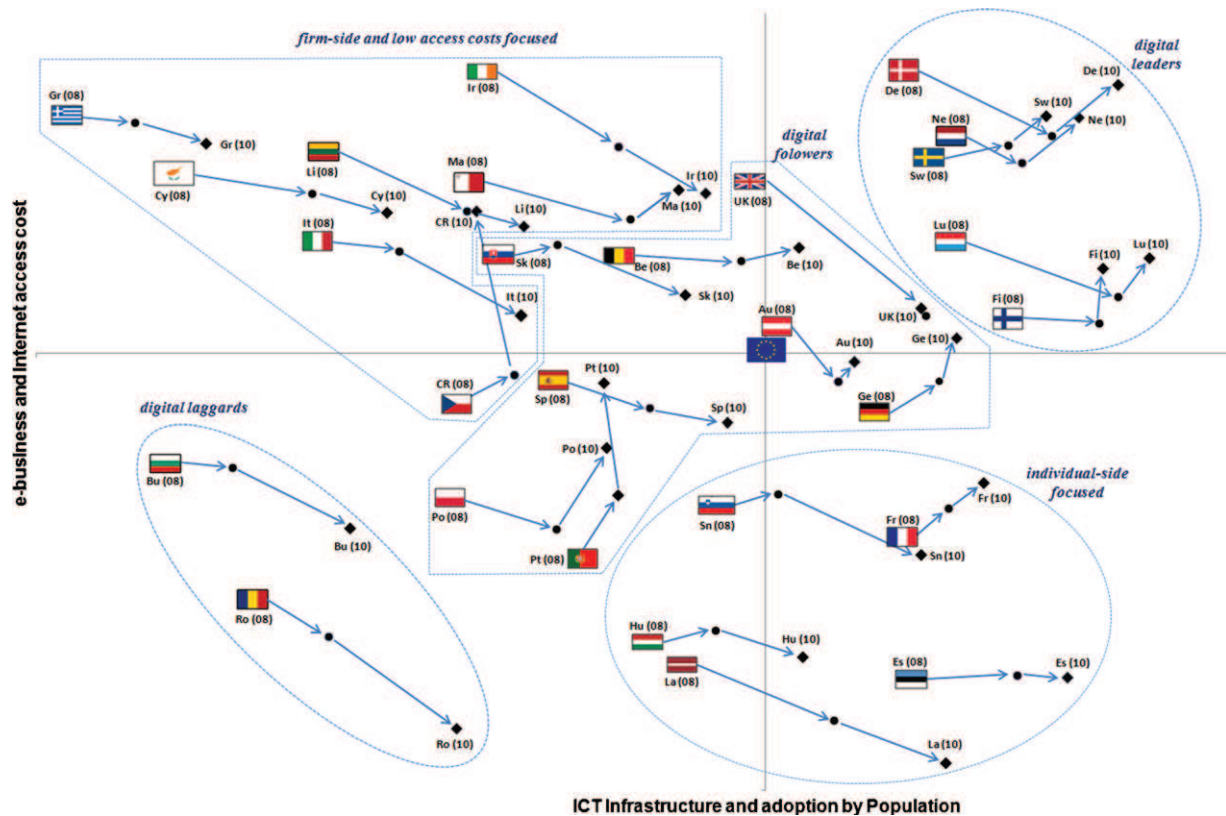**Fig. 5.** Changes in cluster analysis (factors vs. variables).

Fig. 6. Movement of the EU-27 countries in the two dimensions of the digital divide between the years of 2008 and 2010.

information on two of the 16 used. The percentage of individuals using mobile devices to access the Internet (Mob) started to be measured only in 2010, and information on the enterprises having a formally defined ICT security policy with a plan of regular review (eSafeE) is only included in a special issue of the Eurostat related to the use of ICTs by firms, which was produced for 2010. To overcome these missing values without compromising our analysis, we maintained the same values as those obtained for 2010. Hence, a short historical evolution of the European digital divide may be drawn without leading to bias in the conclusions, which would have been likely had we simply compared our results with those of earlier studies, since no other study used exactly the same combination of variables and methodologies, not to mention the 27 European Countries. The time period between 2008 and 2010 comprises the period immediately after to the last enlargement of the EU (2007). With this analysis we are hopeful that we will be able to determine whether the European digital divide has in fact been narrowing or widening, which is not consensual. The path movement of the countries is shown in Fig. 6.

When analysing the evolutions of each country individually, in the *ICT Infrastructure and adoption by Population*, some conclusions may be drawn from our results. Latvia was the country which has experienced the greatest improvement amongst all the EU members. This improvement was possible because Latvia was able to double the percentage of public services available online, which significantly boosted its usage by individuals. Moreover, the general use of ICTs and "advanced services" by individuals has grown considerably more than the average of the other European counterparts. Ireland and Denmark were the second and third countries, respectively, to present higher evolutions for this dimension. On the other hand, the Czech Republic and Portugal were the countries, respectively, that have grown the least. Although there was an improvement in the *ICT Infrastructure and adoption by Population*, this was practically negligible. Hence, all of

the European countries grew in their levels of *ICT Infrastructure and adoption by Population* between the years of 2008 and 2010, which can be noted in the general shift from left to right that all countries experienced. With respect to the *e-business and Internet access costs*, the situation is different. The Czech Republic, Portugal, and Germany were the countries with the most significant improvements in dimension. Within the three years included in our analysis, both the Czech Republic and Portugal witnessed a very significant reduction in the percentage of households without Internet because of the access costs. Thus, the values in 2010 were about a third of those seen in 2008. Moreover, the Czech Republic was also able to significantly improve the percentage of firms selling online (eCom) and using e-government services within the time period under consideration. Countries that need to improve their levels of *e-business and Internet access costs*, i.e. above the horizontal axis, should definitely learn from the Czech Republic and Portugal the appropriate measures and policies to achieve this goal. Although in a more moderate way, Germany is also an example of a country which has successfully boosted its e-commerce levels while at the same time made the access to the Internet less expensive. Contrarily to what happens in the first dimension, several countries have, in fact, decreased their levels of *e-business and Internet access costs*. The United Kingdom, Ireland, and Romania were the countries which have had the most significant decreases. The percentages of firms selling online in the United Kingdom were, in 2008, more than twice what we see for 2010. Moreover, there was a decrease in the percentage of firms using e-government services, while the Internet access costs remained constant. Ireland had exactly the same problems, although in a more moderate way. Finally, Romania had an increase in the percentage of households without Internet because of the access costs, while the use of e-government services by firms also decreased. The percentage of firms selling online showed some growth, but this was insufficient to prevent a general decrease of the *e-business and Internet access costs* dimension.

Some countries significantly increased their values in the two dimensions simultaneously. Finland, France, Germany, the Netherlands, Poland, and Sweden achieved significantly positive evolutions, moving on a diagonal bottom-left to top-right path, and therefore show balanced digital development. Another interesting finding is that countries in the same cluster (for 2010) appear to behave in a very similar way. The "*digital leaders*" cluster was the only one that increased in the two dimensions simultaneously. Thus, the digital divide probably will continue to manifest itself in the near future.

5. Discussion

5.1. Discussion of findings

The digital divide appears to have two independent and latent dimensions, which are the *ICT Infrastructure and adoption by Population* and the *e-business and Internet access costs*. The first dimension expresses the availability and use of ICT infrastructure by the population, which means for particular purposes. The "advanced services" are comprised on this dimension. The second dimension, on the other hand, expresses the commercial use of ICTs (e-business) and the Internet access costs. It is interesting to notice that these two dimensions are related to the sector to which they will be used, i.e. the individuals or the firms. If the behaviour of individuals and firms were directly correlated within each country, as one would expect, our two dimensions would probably be different. The price of Internet access (Cost) also emerges as a key determinant to the digital development, considering that this variable has a high negative influence (loading) on the diffusion of e-commerce, e-safety concerns and e-government indicators, meaning that higher costs are associated with lower levels on *e-business and Internet access costs*, as observed in earlier studies [31,41,57]. In this way it becomes preponderant that public authorities ensure the low prices on Internet access in order to deter the digital divide within the EU-27 [57].

Considering the factor scores of each European country, we have Denmark, the Netherlands and Sweden followed by Luxembourg and Finland as the countries with the highest levels for the two dimensions together, being the most advanced countries when it comes to the digital development. We therefore apply the label "*digital leaders*" to these countries. This situation is not surprising, considering that North European Countries are pioneers in promoting digital development [45]. The fact that some countries present high levels in one dimension and, at the same time, low levels on the other, reveals the imbalances in the digital development process itself. Greece, for instance, has the lowest level for *ICT Infrastructure and adoption by Population* and one of the highest levels for the *e-business and Internet access costs*. Estonia shows the inverse situation. These imbalances threaten the national and European *e-strategies*, because like economic development, digital development must be harmonized and horizontal to all sectors within each country. It is therefore imperative that countries strive to achieve a balanced and homogeneous digital development, i.e. focus both on *ICT infrastructure and adoption by population* and on the diffusion of e-business with low Internet access costs. The results from our factor analysis also tell us that it is not accurate to simply classify a country as digitally developed or not, since the majority of European countries present their own strengths and weaknesses, i.e. imbalances.

With the help of the cluster analysis we were able to identify five different digital profiles within the EU. With only a minor loss of information we are able to analyse the digital disparities not between 27 Countries but between five digital development stages, which greatly improves the ease of our analysis, as intended. Some interesting conclusions may also be drawn from the results of cluster

analysis. Denmark, the Netherlands, Sweden, Luxembourg, and Finland are, from the two cluster analysis, identified as the "*digital leaders*" countries. Therefore, the difference between these countries *vis-à-vis* the remaining European ones is very important. The same fact is found with the less developed countries, Bulgaria and Romania, the "*digital laggards*". The edges of the spectrum are, therefore, well defined. To these two countries, efforts to achieve a digital development should be made toward the direction of the "*digital leaders*" cluster (see Fig. 3), focusing on both dimensions simultaneously. Earlier studies [6] on this subject also pointed to these two countries as the least digitally developed ones in EU.

The cluster analysis was also useful in grouping countries with similar digital imbalances, whether these imbalances are related to the two dimensions or even to the overall 16 variables. When considering the cluster based on factor scores, countries within "*individual-side focused*" and "*firm-side and low access costs focused*" clusters must strive for homogenous digital development following their respective arrows (see Fig. 3), in order to achieve the objectives expressed in the Digital Agenda for Europe, i.e. "*individual-side focused*" countries need to move toward "*digital leaders*" countries in a bottom to top vertical direction, emphasizing the development of *e-business and Internet access*, while "*firm-side and low access costs focused*" countries need to move in a left to right horizontal direction focusing on the *ICT Infrastructure and adoption by Population*. With this proposed movement we are not arguing that these countries should not try to improve the other dimension as well, but for a question of homogeneity between enterprises and individuals, if they had to choose one dimension only for improvement, they should focus on the one indicated as the more vulnerable at the time. Hence, unlike Bulgaria and Romania, these countries do not need to focus on the two digital dimensions simultaneously, which is an advantage. It is a priority that the national leaders of these imbalanced, or developing, countries can learn from their northern European counterparts so that the accurate measures can be taken in order to bridge this gap.

As we sought to arrive at some possible explanations for this European digital divide, the entrance year to the Union appears to have a word to say. Digitally speaking, the integration process within the EU is not yet completed, especially for the countries that entered in 2004 or thereafter. School attendance has a marginal effect on the digital divide, which contrasts with the views of some authors who hypothesized that this would be a significant factor [7,8,77,78]. This factor may be related to the recent entrance of eastern European countries that already have highly developed education systems with not so significant digital developments. Thus, we suspect that the use of ICT in these countries will accelerate soon, when the impact of the European structural funds, addressed to the information society programmes, strengthens. When that happens, these countries may show new infrastructures and highly educated individuals, who will be more likely to use them. As suggested in the literature, digital imbalances are directly related to economic wealth, since the GDP per Capita has been proven to be very significant, at a confidence level higher than 99.9%. The EU can be considered especially vulnerable to these two factors, considering that unlike the US, it is presently a mix of 27 different countries, with many profound differences amongst them.

Finally the last step in our analysis of the European digital divide was to determine the paths of the European countries with regard to the two dimensions found, from 2008 to 2010. This movement is in respect to the two dimensions because the comparison between the two cluster analyses gave us the assurance that the two dimensions extracted are reliable in comparison to the whole dataset. The movement of the EU countries revealed a surprising fact: The clusters found for the year of 2010 appear to exist, at least since 2008, with exception to the frontier countries between "*firm-side and low access costs focused*"

and "*digital followers*" clusters, emphasizing the reliability of our results. More importantly, the "*digital leaders*" countries even in 2008 showed significantly higher levels for the two dimensions than the rest of the EU in 2010. Hence, the "*digital leaders*" have, at the minimum, three years of advance in terms of digital development when compared to the EU average (note that the EU flag in Fig. 6 represent the average of the two factors for 2010). The digital divide is, in fact, rooted within the EU. The only cluster that has overall improved the levels of *e-business and Internet access costs* was the one containing the "*digital leaders*". Coincidently, these countries are also the ones with higher levels of *ICT Infrastructure and adoption by Population*. One may ask, as a result, whether this fact means that it is only from a certain point of ICT infrastructures availability and individual adoption that e-business is effectively triggered? Finally, one last controversial [79] question, and perhaps the most important, must be asked: Is the European digital divide narrowing or widening? We argue that both processes are, in fact, taking place. That is, in one dimension – *ICT Infrastructure and adoption by Population* – there is evidence that this divide may be narrowing. Although all clusters in our analysis showed growth in this dimension, the "*digital laggards*", and those with digital imbalances were the ones that increased the most. Therefore, this gap is narrowing. On the other hand, the already "*digital leaders*" were the ones presenting positive developments in the second dimension – *e-business and Internet access costs*. Thus, the digital asymmetries between countries in this particular aspect of the divide are widening. Perhaps the financial and economic crisis of 2008, which originated in the US, has influenced the increasing percentage of households without Internet because of the access costs, and the reduction of e-commerce levels in some of these countries, which justifies in some way the fact that several countries showed decreasing levels of the *e-business and Internet access costs*.

Nevertheless, as pointed by Dimaggio [80], initiatives to minor digital inequalities have emphasized mainly the access to technologies, which may not be sufficient. Above all it is necessary to combine the efforts of public authorities, private organizations, and the population itself to bridge this divide [29], since a single community or sector cannot do it alone [41].

5.2. Limitations of the study

In spite of our effort to offer a complete and multidimensional analysis, some limitations must be recognized. First, our empirical application consists of just 16 variables, and, some features of the information society may not be covered. Second, we analysed the digital divide within the EU, which means that all indicators used were concerned with aggregate national realities, meaning that internal, domestic digital divide gaps may not be covered. The third limitation is related to the lack of available data for the years of 2008 and 2009 for two of the 16 variables considered. As we did not wish to make any assumptions about its evolution, we had to use the levels of 2010. Finally, in fourth place, our analysis refers to the digital divide at a specific point in time, the years between 2008 and 2010. Changes in this context are likely to occur rapidly, and our findings may soon become outdated.

6. Conclusions

Based on multivariate statistical methods, we analyse the digital divide within the EU-27. We find that, in fact, a digital divide still exists within the EU, despite all the investments and policies to narrow it in recent years. Thus, the countries identified as the least digitally developed in earlier studies remain basically the same, whereas the same fact is obtained for the most digitally developed countries. The digital development has two independent dimensions, and we detect five digital profiles amongst the 27 Member States. The digital disparities are correlated by economic asymmetries between the countries, while the entrance year also appears to influence the divides. On the other hand, the school attendance of the population does not appear to have a significant importance on the digital divide, which goes against what some studies have reported in the past. Finally, we also concluded, based on the path movement of each country between 2008 and 2010 that in one dimension there is evidence that the European digital divide is narrowing, while in the other it appears to be widening.

Acknowledgement

Appendix A
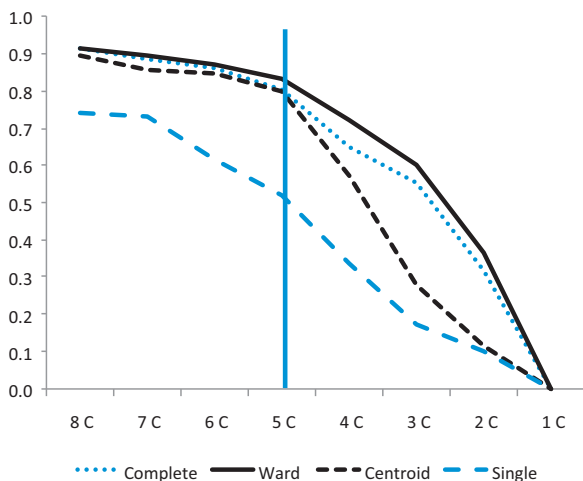
See Figs. 7 and 8 and Table 6.

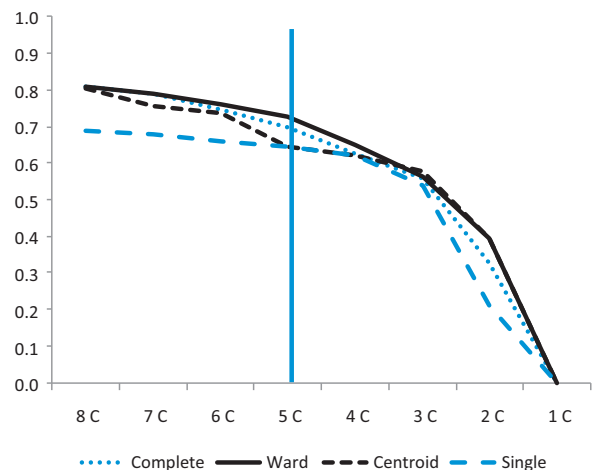**Fig. 7.** *R*-square plot for cluster analysis based on factor scores.

Fig. 8. *R*-square plot for cluster analysis based on original 16 variables.

Table 6
Descriptive statistics for the identified clusters (variables).

	Digital laggards		Individual-side		Firm-side and costs		Digital followers		Digital leaders		Kruskal–Wallis
	Average	St Dev	Average	St Dev	Average	St Dev	Average	St Dev	Average	St Dev	*p*-Value
HsInt	37.50	6.4	66.00	6.0	58.22	6.1	75.00	4.8	87.20	4.0	<0.0001
BroRt	13.80	0.1	23.84	5.0	19.46	2.8	27.80	3.7	34.18	4.1	0.0003
IntPop	38.00	5.7	66.80	6.0	54.22	9.2	70.33	7.6	86.20	2.0	<0.0001
Mob	6.50	2.1	24.20	8.2	17.67	8.5	31.50	7.9	40.80	11.2	0.0002
IntSrc	26.00	0.0	59.00	4.0	46.44	9.0	60.67	6.8	78.80	3.3	<0.0001
Cost	15.50	9.2	15.60	4.5	7.00	2.1	3.83	1.7	1.80	2.2	0.0123
eBank	2.50	0.7	42.60	18.5	22.78	9.2	41.50	6.1	71.00	8.7	<0.0001
eLearn	16.50	4.9	41.60	8.8	31.00	7.5	40.17	3.4	58.80	14.5	0.0008
email	33.00	2.8	61.20	6.7	49.11	11.1	66.00	8.3	82.80	3.6	<0.0001
eHealth	16.00	4.2	37.40	4.5	26.89	5.6	35.83	7.0	51.40	7.2	0.0002
eGovI	11.00	5.7	36.80	7.9	22.44	7.1	33.83	5.6	61.20	6.5	<0.0001
eGovE	57.00	9.9	77.80	6.9	82.00	9.2	75.00	7.5	92.60	2.8	0.0054
eGovS	65.00	7.1	86.57	12.3	76.02	19.3	95.29	8.4	91.37	10.9	0.0560
eCom	5.00	1.4	9.20	2.3	12.00	6.6	19.00	4.7	20.20	4.9	0.0304
Serv	56.54	23.7	276.15	111.0	285.37	293.4	997.63	342.3	1613.42	446.8	0.0007
eSafeE	8.00	1.4	14.60	5.0	28.00	9.1	27.83	2.1	36.60	8.1	0.0369

References

[1] WSIS, Tunis commitment, Second Phase of the World Summit on the Information Society, Tunis, 2005, p. 5.
[2] WSIS, World Summit on the Information Society: declaration of principles, World Summit on the Information Society, Geneva, 2003, p. 9.
[3] European Commission, Europe 2020 – A Strategy for Smart, Sustainable and Inclusive Growth, Brussels, 2010, p. 37.
[4] European Commission, A Digital Agenda for Europe, Brussels, 2010, p. 42.
[5] European Commission, Investing Today for Tomorrow's Europe (COM/2009/36/final), European Commission, Brussels, 2009.
[6] M.R. Vicente, F. Gil-de-Bernabé, Assessing the broadband gap: from the penetration divide to the quality divide, Technological Forecasting and Social Change 77, 2010, pp. 816–822.
[7] M. Billon, R. Ezcurra, F. Lera-López, Spatial effects in website adoption by firms in european regions, Growth & Change 40, 2009, pp. 54–84.
[8] M. Billon, R. Ezcurra, F. Lera-López, The spatial distribution of the internet in the european union: does geographical proximity matter? European Planning Studies 16, 2008, pp. 119–142.
[9] D.J. Gunkel, Second thoughts: toward a critique of the digital divide, New Media & Society 5, 2003, pp. 499–522.
[10] N.-G. Dragulanescu, Social Impact of the "Digital Divide" in a Central-Eastern European Country, The International Information & Library Review 34, 2002, pp. 139–151.
[11] P.B. Brandtzæg, J. Heim, A. Karahasanovic, Understanding the new digital divide – a typology of Internet users in Europe, International Journal of Human–Computer Studies 69, 2011, pp. 123–138.
[12] OECD Understanding the Digital Divide, OECD Publications, Paris, 2001, p. 32.
[13] J.J.P.-A. Hsieh, A. Rai, M. Keil, Understanding digital inequality: comparing continued use behavioral models of the social-economically advantaged and disavantaged, MIS Quarterly 32, 2008, pp. 97–126.
[14] S. Okazaki, What do we know about mobile Internet adopters? A cluster analysis Information & Management 43, 2006, pp. 127–141.
[15] M. Warschauer, Reconceptualizing the digital divide, First Monday, 2002.
[16] J.C. Bertot, The multiple dimensions of the digital divide: more than the technology 'haves' and 'haves nots', Government Information Quarterly 20, 2003, pp. 185–191.
[17] H. Ono, M. Zavodny, Digital inequality: a five country comparison using microdata, Social Science Research 36, 2007, pp. 1135–1155.
[18] I.f.S. Unesco, Measuring and monitoring the information and knowledge societies: a statistical challenge, 2003.
[19] US Department of Commerce, Falling through the Net I: A Survey of the 'Have-Nots' in Urban and Rural America, US Department of Commerce, Washington, DC, 1995.
[20] US Department of Commerce, Falling through the Net II: New Data on the Digital Divide, US Department of Commerce, Washington, DC, 1998.
[21] US Department of Commerce, Falling through the Net III: Defining the Digital Divide, US Department of Commerce, Washington, DC, 1999.
[22] US Department of Commerce, Falling through the Net IV: Toward Digital Inclusion, US Department of Commerce – Economic and Statistics Administration, Washington, DC, 2000.
[23] M.R. Vicente, A.J. Lopez, A multidimensional analysis of the disability digital divide: some evidence for internet use, The Information Society: An International Journal 26, 2010, pp. 48–64.
[24] US Department of Commerce, A Nation Online: How Americans Are Expanding Their Use of the Internet, US Department of Commerce – Economic and Statistics Administration, Washington, DC, 2002.
[25] M.R. Vicente, A.J. López, Patterns of ICT diffusion across the European Union, Economics Letters 93, 2006, pp. 45–51.

[26] M.R. Vicente, A.J. Lopez, Some empirical evidence on Internet diffusion in the New Member States and Candidate Countries of the European Union, Applied Economics Letters 15, 2008, pp. 1015–1018.
[27] M.R. Vicente, A.J. Lopez, What drives broadband diffusion? Evidence from Eastern Europe Applied Economics Letters 17, 2010, pp. 51–54.
[28] E.M. Crenshaw, K.K. Robison, Globalization and the digital divide: the roles of structural conduciveness and global connection in internet diffusion, Social Science Quarterly (Blackwell Publishing Limited) 87, 2006, pp. 190–207.
[29] S. Dewan, F.J. Riggins, The digital divide: current and future research directions, Journal of the Association for Information Systems 6, 2005, pp. 298–337.
[30] OECD, Guide to Measuring the Information Society, Secretary-General of the OECD, Paris, 2009.
[31] M.R.V. Cuervo, A.J.L. Menéndez, A multivariate framework for the analysis of the digital divide: Evidence for the European Union-15, Information & Management 43, 2006, pp. 756–766.
[32] The World Bank, Information and Communications for Development 2009: Extending Reach and Increasing Impact, Washington, DC, 2009.
[33] OECD, The economic impact of ICT: measurement, evidence and implications, in: Organization for Economic Cooperation and Development (Ed.), Paris, 2004.
[34] S. Wattal, D. Schuff, M. Mandviwalla, C.B. Williams, Web 2.0 and politics: the 2008 U.S. Presidential Election and an e-Politics Research Agenda, MIS Quarterly 34, 2010, pp. 669–688.
[35] B. Carlsson, The digital economy: what is new and what is not? Structural Change and Economic Dynamics 15, 2004, pp. 245–264.
[36] J. Jalava, M. Pohjola, The roles of electricity and ICT in economic growth: case Finland, Explorations in Economic History 45, 2008, pp. 270–287.
[37] Ç.A. Çilan, B.A. Bolat, E. Coskun, Analyzing digital divide within and between member and candidate countries of European Union, Government Information Quarterly 26, 2009, pp. 98–105.
[38] N.S.K. Facer, Beyond the digital divide: rethinking digital inclusion for the 21st century 2007, p. 40.
[39] European Commission, Bridging the Broadband Gap, Brussels, 2006, p. 10.
[40] S. Brooks, P. Donovan, C. Rumble, Developing nations, the digital divide and research databases, Serials Review 31, 2005, pp. 270–278.
[41] T. Unwin, G. de Bastion, Digital divide, in: K. Rob, T. Nigel (Eds.), International Encyclopedia of Human Geography, Elsevier, Oxford, 2009, pp. 191–197.
[42] The World Bank, Information and Communications for Development: Global Trends and Policies, The World Bank, Washington, DC, 2006.
[43] A.R. Bakker, Health care and ICT, partnership is a must, International Journal of Medical Informatics 66, 2002, pp. 51–57.
[44] M. Cukusic, N. Alfirevic, A. Granic, Z. Garaca, e-Learning process management and the e-learning performance: results of a European empirical study, Computers & Education 55, 2010, pp. 554–565.
[45] M. Castells, P. Himanen, The Information Society and the Welfare State: The Finnish Model, 1st ed., Oxford University Press, New York, 2002.
[46] M. Castells, The Network Society, vol. III, 1st ed., Fundação Calouste Gulbenkian, Lisbon, 2007.
[47] V. Vehovar, P. Sicherl, T. Husing, V. Dolnicar, Methodological challenges of digital divide measurements, The Information Society 22, 2006, pp. 279–290.
[48] S.-Y.T. Lee, R. Gholami, T.Y. Tong, Time series analysis in the assessment of ICT impact at the aggregate level – lessons and implications for the new economy, Information & Management 42, 2005, pp. 1009–1022.
[49] C.-H. Wang, Y. McLee, J.-H. Kuo, Mapping the intellectual structure of digital divide, International Journal of Social Science and Humanity 1, 2011, pp. 49–54.
[50] G. Sciadas, P. Guiguère, L. Adarn, From the Digital Divide to Digital Opportunities: Measuring Infostates for Development, Orbicom International Secretariat, Geneva, Switzerland, 2005.
[51] A. Emrouznejad, E. Cabanda, R. Gholami, An alternative measure of the ICT-opportunity Index, Information & Management 47, 2010, pp. 246–254.

[52] OECD, Handbook on Constructing Composite Indicators: Methodology and User Guide, Joint Research Center (JRC) of the European Commission, Organisation for Economic Co-operation and Development, OECD, Paris, 2008.

[53] G. Bruno, E. Esposito, A. Genovese, K.L. Gwebu, A critical analysis of current indexes for digital divide measurement, The Information Society 27, 2010, pp. 16–28.

[54] M.R. Vicente, A.J. López, Assessing the regional digital divide across the European Union-27, Telecommunications Policy 35, 2010, pp. 220–237.

[55] S.H. Doong, S.-C. Ho, The impact of ICT development on the global digital divide, Electronic Commerce Research and Applications 11, 2012, pp. 518–533.

[56] S. Dewan, D. Ganley, K.L. Kraemer, Across the digital divide: a cross-country multi-technology analysis of the determinants of IT penetration, Journal of the Association for Information Systems 6, 2005, pp. 409–431.

[57] S. Dewan, D. Ganley, K.L. Kraemer, Complementarities in the diffusion of personal computers and the internet: implications for the global digital divide, Information Systems Research 2009, pp. 1–17.

[58] M.D. Chinn, R.W. Fairlie, The determinants of the global digital divide: a cross-country analysis of computer and internet penetration, Oxford Economic Papers – New Series 59, 2007, pp. 16–44.

[59] E. Hargittai, Weaving the Western Web: explaining differences in Internet connectivity among OECD countries, Telecommunications Policy 23, 1999, pp. 701–718.

[60] S. Kiiski, M. Pohjola, Cross-country diffusion of the Internet, Information Economics and Policy 14, 2002, pp. 297–310.

[61] M. Pohjola, The Adoption and Diffusion of ICT Across Countries: Patterns and Determinants. The New Economy Handbook, Academic Press, 2003 pp. 230–247.

[62] M.D. Chinn, R.W. Fairlie, ICT use in the developing world: an analysis of differences in computer and internet penetration, Review of International Economics 18, 2010, pp. 153–167.

[63] M. Billon, R. Marco, F. Lera-Lopez, Disparities in ICT adoption: a multidimensional approach to study the cross-country digital divide, Telecommunications Policy 33, 2009, pp. 596–610.

[64] R.J. Kauffman, A.A. Techatassanasoontorn, Is there a global digital divide for digital wireless phone technologies? Journal of the Association for Information Systems 6, 2005, pp. 338–381.

[65] E. Ferro, N.C. Helbig, J.R. Gil-Garcia, The role of IT literacy in defining digital divide policy needs, Government Information Quarterly 28, 2011, pp. 3–10.

[66] U. Bunz, C. Curry, W. Voon, Perceived versus actual computer–email–web fluency, Computers in Human Behavior 23, 2007, pp. 2321–2344.

[67] S.-U. Guan, Y. Yang, SAFE: secure agent roaming for e-commerce, Computers & Industrial Engineering 42, 2002, pp. 481–493.

[68] G. Dhillon, G. Torkzadeh, Value-focused assessment of information system security in organizations, Information Systems Journal 16, 2006, pp. 293–314.

[69] C. Kim, W. Tao, N. Shin, K.-S. Kim, An empirical study of customers perceptions of security and trust in e-payment systems, Electronic Commerce Research and Applications 9, 2010, pp. 84–95.

[70] J. Spicer, Making Sense of Multivariate Data Analysis: An Intuitive Approach, Sage Publications, Thousand Oaks, 2005.

[71] J.F. Hair, R.E. Anderson, R.L. Tatham, W.C. Black, Multivariate Data Analysis: With Readings, Prentice Hall International, London, 1995.

[72] J. Lee, H.J. Kim, M.J. Ahn, The willingness of e-Government service adoption by business users: the role of offline service quality and trust in technology, Government Information Quarterly 28, 2011, pp. 222–230.

[73] I. Jolliffe, Principal component analysis, Encyclopedia of Statistics in Behavioral Science, John Wiley & Sons, Ltd, 2005.

[74] P.R. Peres-Neto, D.A. Jackson, K.M. Somers, How many principal components? Stopping rules for determining the number of non-trivial axes revisited Computational Statistics & Data Analysis 49, 2005, pp. 974–997.

[75] S. Sharma, Applied Multivariate Techniques, John Wiley & Sons, Inc., New York, 1996.

[76] F. Leisch, A toolbox for *K*-centroids cluster analysis, Computational Statistics & Data Analysis 51, 2006, pp. 526–544.

[77] M.G. Quibria, S.N. Ahmed, T. Tschang, M.-L. Reyes-Macasaquit, Digital divide: determinants and policies with special reference to Asia, Journal of Asian Economics 13, 2003, pp. 811–825.

[78] A. Goldfarb, The (teaching) role of universities in the diffusion of the Internet, International Journal of Industrial Organization 24, 2006, pp. 203–225.

[79] Dan M. Grigorovici, Jorge R. Schement, R.D. Taylor, Weighing the intangible: towards a theory-based framework for information society indices, in: E. Bohlin, S. Levin, N. Sung, C.-H. Yoon (Eds.), Global Economy and Digital Society, Elsevier Science, 2004, p. 31.

[80] P. DiMaggio, E. Hargittai, C. Celeste, S. Shafer, Digital inequality: from unequal access to differentiated use, in: K. Neckerman (Ed.), Digital Inequality, Russell Sage Foundation, New York, 2004.

Frederico Cruz-Jesus is currently a Ph.D. student at Instituto Superior de Estatística e Gestão de Informação, Universidade Nova de Lisboa. He received his Masters degree in Statistics and Information Management, with specialization in Business Intelligence and Knowledge Management from the same University (2012). His research interests currently include the issue of the digital divide, the use and value of business analytics, and the adoption of technological innovations by individuals and firms. Professionally, Frederico Cruz-Jesus has worked in major retail and banking Portuguese firms. Currently he is in the insurance sector, in the area of CRM at Allianz Portugal.

Tiago Oliveira is Invited Assistant Professor at ISEGI, Universidade Nova de Lisboa. He holds a Ph.D. from the Universidade Nova de Lisboa in Information Management. His research interests include technology adoption, digital divide and privacy. He has published papers in several academic journals and conferences, including the Industrial Management & Data Systems, Applied Economics Letters, Electronic Journal of Information Systems Evaluation, and American Journal of Mathematical and Management Sciences among others. Additional detail can be found in http://www.isegi.unl.pt/toliveira/.

Fernando Bacao is currently Associate Professor at the School of Statistics and Information Management, Universidade Nova de Lisboa. He holds a Ph.D. in Information Management and his research interests include decision support systems, with particular emphasis on geospatial applications of neural networks. Recently, he has extended his research interests around the digital divide and the role of business analytics in leveraging information for innovation and competitive advantage. His research work has appeared in journals such as Soft Computing, Landscape Ecology, International Journal of Wildland Fire, Environment and Planning B, International Journal of Geographical Information Science, Computers and Geosciences. Additional detail can be found in http://www.isegi.unl.pt/fbacao/.